

# ISOMETRIES OF INFINITE DIMENSIONAL HILBERT GEOMETRIES

BAS LEMMENS, MARK ROELANDS, AND MARTEN WORTEL

**ABSTRACT.** In this paper we extend results by De la Harpe concerning the isometries of strictly convex Hilbert geometries, and the characterisation of the isometry groups of Hilbert geometries on finite dimensional simplices, to infinite dimensions. The proofs rely on a mix of geometric and functional analytic methods.

## 1. INTRODUCTION

In [9] Hilbert introduced a collection of metric spaces that are natural deformations of finite dimensional real hyperbolic spaces. Although Hilbert limited his construction to finite dimensional spaces, it has a straightforward extension to infinite dimensions. Indeed, let  $\Omega$  be a convex subset of a (not necessarily finite-dimensional) vector space  $Y$  and suppose that for each  $x \neq y \in \Omega$ , the straight line  $\ell_{xy}$  through  $x$  and  $y$  has the property that  $\Omega \cap \ell_{xy}$  is an open and bounded line segment in  $\ell_{xy}$ . In that case one can define Hilbert's metric on  $\Omega$  as follows. For  $x \neq y$  in  $\Omega$ , let  $x'$  and  $y'$  be the end-points of the segment  $\ell_{xy} \cap \Omega$  such that  $x$  is between  $x'$  and  $y$ , and  $y$  is between  $y'$  and  $x$ . Now *Hilbert's metric on  $\Omega$*  is given by

$$\delta_H(x, y) := \log[x', x, y, y'] \quad \text{for } x \neq y \text{ in } \Omega,$$

where

$$[x', x, y, y'] := \frac{|x' - y|}{|x' - x|} \frac{|y' - x|}{|y' - y|}$$

is the *cross-ratio*, and  $\delta_H(x, x) = 0$  for all  $x \in \Omega$ . The metric space  $(\Omega, \delta_H)$  is usually called the *Hilbert geometry on  $\Omega$* . In particular, the open unit ball in an infinite dimensional separable Hilbert space equipped with  $\frac{1}{2}\delta_H$ , is precisely Klein's model of the infinite dimensional hyperbolic space  $\mathbf{H}^\infty$  studied in [4]. A recent extensive overview of the theory of Hilbert geometries can be found in [7].

It is well known that Hilbert's metric has important applications in the analysis of linear, and nonlinear, operators on cones both in finite and infinite dimensions, see [11, 12, 17]. In mathematical analysis one often works with

---

*Date:* May 19, 2014.

2010 *Mathematics Subject Classification.* Primary 58B20; Secondary 22F50, 46B04.

Bas Lemmens was supported by EPSRC grant EP/J008508/1.

Mark Roelands was supported by EPSRC grant EP/J500446/1.

Marten Wortel was supported by EPSRC grant EP/J008508/1.

Birkhoff's version of Hilbert's metric, which provides a slightly more general set up than the one outlined above. Suppose  $\Omega$  is a convex and bounded subset in a normed space  $Y$ . Let  $X := Y \oplus \mathbb{R}$  and  $\varphi: X \rightarrow \mathbb{R}$  be given by  $\varphi(x, \lambda) := \lambda$ . Then we can define the cone

$$C_\Omega := \{\lambda(x, 1) : \lambda \geq 0 \text{ and } x \in \overline{\Omega}\} \subset X,$$

where we can recover  $\Omega$  as the cross section  $\Sigma_\varphi := \{x \in C_\Omega : \varphi(x) = 1\}$ . The advantage of using cones is that Hilbert's metric can be defined in terms of the partial order induced by the cone (Birkhoff's version), and so we can use methods from the theory of partially ordered vector spaces.

However, the assumption that  $\Omega$  is bounded is an unnecessary restriction. Indeed, as an example one can consider the set

$$\Omega := \left\{ f \in C[0, 1] : \int_0^1 f(t) dt = 1, \quad f([0, 1]) \subset (0, \infty) \right\},$$

which is unbounded in the space of continuous functions  $C[0, 1]$  with the supremum norm, but  $\Omega$  does admit Hilbert's metric. In Section 2 we will discuss, beside the relevant definitions, the precise relation between Hilbert geometries and cones. In fact, we will show the following theorem.

**Theorem 1.1.** *If  $Y$  is a vector space and  $\Omega \subset Y$  is a convex set on which  $\delta_H$  is well defined, then there exists an order unit space  $(X, C, u)$  and a strictly positive state  $\varphi$  on  $X$  such that  $\Omega$  is affinely isomorphic to  $\Sigma_\varphi$ .*

*Conversely, if  $(X, C, u)$  is an order unit space with a strictly positive state  $\varphi$ , then  $\Sigma_\varphi$  is a convex set on which  $\delta_H$  is well defined.*

The main goal of this paper is to take up a suggestion by De la Harpe [10, §3.3] and study the isometries of infinite dimensional Hilbert geometries. We shall extend two results from [10] to infinite dimensions. Firstly, we prove in Section 3 the following theorem, which generalises [10, Proposition 3].

**Theorem 1.2.** *If  $\Omega_1$  and  $\Omega_2$  are strictly convex Hilbert geometries, and  $f: \Omega_1 \rightarrow \Omega_2$  is an isometry of  $(\Omega_1, \delta_H)$  into  $(\Omega_2, \delta_H)$ , then  $f$  is a projective linear homomorphism.*

In this theorem,  $f$  is projective linear with respect to the space  $X$  from Theorem 1.1.

The proof of Theorem 1.2 will follow a similar approach as the one in taken in [10]. However, as we are working in infinite dimensions, there need not be a basis, so no use can be made of the fundamental theorem of projective geometry. Instead we establish an extension result, Proposition 3.3, which can be combined with Zorn's Lemma to prove Theorem 1.2.

The second result concerns the isometries of Hilbert geometries on infinite dimensional simplices. Recall that any two  $n$ -dimensional simplices are affinely isomorphic, and hence they are isometrically isomorphic as Hilbert geometries. Consider the following  $n$ -simplex:

$$\Delta_n^\circ := \{x \in \mathbb{R}^{n+1} : x_i > 0 \text{ for all } i \in \{1, \dots, n+1\} \text{ and } \sum_{i=1}^{n+1} 2^{-i} x_i = 1\}.$$

A natural generalisation is to consider infinite sequences. Theorem 1.1 shows that we need an order unit. The constant one sequence is a natural choice for an order unit, which yields the sequence space  $\ell^\infty$  and the “infinite dimensional simplex”

$$\Delta_\infty^\circ := \{x \in \ell^\infty : x_i > 0 \text{ for all } i \in \mathbb{N} \text{ and } \varphi(x) := \sum_{i=1}^\infty 2^{-i} x_i = 1\}.$$

As  $\ell^\infty$  is an abstract  $M$ -space with an order unit, it follows from Kakutani’s representation theorem that  $\ell^\infty$  is isometrically order-isomorphic to  $C(K)$  for some compact Hausdorff space  $K$ , see [15, Theorem 1.b.6]. Under this isomorphism the state  $\varphi$  turns into a strictly positive Borel measure  $\mu$  on  $K$ , i.e.,  $\int f d\mu > 0$  for all  $f \in C(K)$  with  $f \neq 0$  and  $f(x) \geq 0$  for all  $x \in K$ . So,  $\Delta_\infty^\circ$  is a special case of

$$\Delta(K, \mu)^\circ := \{f \in C(K) : f(x) > 0 \text{ for all } x \in K \text{ and } \int f d\mu = 1\}.$$

For the Hilbert geometries  $(\Delta(K, \mu)^\circ, \delta_H)$  we have the following result.

**Theorem 1.3.** *If  $K_1$  and  $K_2$  are compact Hausdorff spaces with strictly positive Borel measures  $\mu_1$  and  $\mu_2$ , respectively, then  $h : \Delta(K_1, \mu_1)^\circ \rightarrow \Delta(K_2, \mu_2)^\circ$  is a surjective Hilbert’s metric isometry if and only if there exist  $\varepsilon \in \{-1, 1\}$ , a homeomorphism  $\vartheta : K_2 \rightarrow K_1$ , and  $g \in C(K_2)$  with  $g(x) > 0$  for all  $x \in K_2$  such that*

$$h(f) = \frac{g(f \circ \vartheta)^\varepsilon}{\int g(f \circ \vartheta)^\varepsilon d\mu_2}.$$

*If  $(K_1, \mu_1) = (K_2, \mu_2) = (K, \mu)$  and  $|K| \geq 3$ , then the isometry group is given by*

$$\text{Isom}(\Delta(K, \mu)^\circ, \delta_H) \cong \overline{C(K)} \rtimes (C_2 \times \text{Homeo}(K)),$$

*where  $C_2$  is the cyclic group of order 2.*

Theorem 1.3 generalises the characterisation of the isometries of Hilbert geometry on finite dimensional simplices obtained in [10], see also [6, Theorem 5.1] and [14, Theorem 1.2]. Our proof is similar in spirit as the proof of the Banach-Stone theorem; cf. [5].

The isometries of Hilbert geometries have been studied for numerous domains. In particular, Lemmens and Walsh studied the isometry groups of polytopal Hilbert geometries in [14]. Bosch  [3] has characterised the Hilbert’s metric isometry groups for projective domains of symmetric cones. Recently Walsh [19] showed that every isometry of a finite dimensional Hilbert geometry is a projective linear automorphism except when the domain comes from a non-Lorentzian symmetric cone. The group of Hilbert’s metric isometries on the projective domain of the cone of positive self-adjoint operators on a complex Hilbert space was determined by Moln r in [16].

## 2. GENERAL HILBERT GEOMETRIES

In this section we introduce the basic concepts and provide a proof for Theorem 1.1.

A cone  $C$  in a vector space  $X$  is a convex set such that  $\lambda C \subseteq C$  for all  $\lambda \geq 0$  and  $C \cap (-C) = \{0\}$ . A cone  $C$  induces a partial ordering  $\leq_C$  on  $C$  by  $x \leq_C y$  if  $y - x \in C$ . The cone is said to be *Archimedean* if for each  $x \in X$  and  $y \in C$  with  $nx \leq_C y$  for all  $n = 1, 2, 3, \dots$  we have that  $x \leq_C 0$ . An element  $u \in C$  is called an *order unit* if for each  $x \in X$  there exists  $\lambda > 0$  such that  $x \leq_C \lambda u$ . The triple  $(X, C, u)$  is called an *order unit space*, if  $C$  is an Archimedean cone in  $X$  and  $u$  is an order unit for  $C$ .

Given an order unit space  $(X, C, u)$ , the space  $X$  can be equipped with the so-called *order unit norm*,

$$\|x\|_u := \inf\{\lambda > 0: -\lambda u \leq_C x \leq_C \lambda u\}.$$

By [1, Theorem 2.55(2)],  $C$  is closed under  $\|\cdot\|_u$ . Note also that  $\|\cdot\|_u$  is a *monotone norm*, i.e.,  $\|x\|_u \leq \|y\|_u$  for all  $0 \leq_C x \leq_C y$ , and hence  $C$  is normal with respect to  $\|\cdot\|_u$ . Recall that a cone  $C$  is called *normal* in a normed space  $(X, \|\cdot\|)$  if there exists a constant  $\kappa > 0$  such that  $\|x\| \leq \kappa\|y\|$  whenever  $0 \leq_C x \leq_C y$ . Furthermore,  $C$  has nonempty interior,  $C^\circ$ , with respect to  $\|\cdot\|_u$ , as the following lemma shows.

**Lemma 2.1.** *If  $(X, C, u)$  is an order unit space, then the set of order units of  $C$  coincides with  $C^\circ$ .*

*Proof.* Every interior point is an order unit by [1, Lemma 2.5]. Conversely, if  $x \in C$  is an order unit, then there exists  $M > 0$  such that  $u/M \leq x$ . If  $\|y\|_u \leq 1/M$  then  $x - y \geq x - u/M \geq 0$ , so  $x \in C^\circ$ .  $\square$

Throughout the paper we shall always assume that an order unit space is equipped with the order unit norm.

A linear functional  $\varphi: X \rightarrow \mathbb{R}$  on an order unit space  $(X, C, u)$  is said to be *strictly positive* if  $\varphi(C \setminus \{0\}) \subseteq (0, \infty)$ . A linear functional  $\varphi$  with  $\varphi(u) = 1$  is called a *state* of  $(X, C, u)$ . Strictly positive states are always continuous and  $\|\varphi\| = 1$ , as can be seen from the following lemma.

**Lemma 2.2.** *Let  $(X_1, C_1, u_1)$  and  $(X_2, C_2, u_2)$  be order unit spaces. If  $T: X_1 \rightarrow X_2$  is a linear map such that  $T(C_1) \subseteq C_2$ , then  $T$  is continuous with  $\|T\| = \|Tu_1\|_{u_2}$ .*

*Proof.* If  $\|x\|_{u_1} \leq 1$ , then  $-u_1 \leq_{C_1} x \leq_{C_1} u_1$ , so that  $-Tu_1 \leq_{C_2} Tx \leq_{C_2} Tu_1$ . The statement now follows from the monotonicity of  $\|\cdot\|_{u_2}$ .  $\square$

*Remark 2.3.* We note that if  $(X, C, u)$  is an order unit space, there may not exist a strictly positive functional. Indeed, if  $X$  is the vector space of bounded functions on an uncountable set and  $C$  is the cone of functions taking nonnegative values everywhere, then no strictly positive functional exists, see [1, Exercise 6, Section 1.7]. However, if  $X$  is separable, then a strictly positive state always exists. Indeed, in that case the unit ball  $B_{X^*}$  of  $X^*$  (which is weak\*-compact) is weak\*-metrizable, and so there exists a sequence of states  $\varphi_n$  which is weak\*-dense in the set of all states. A standard argument then shows that the state  $\varphi := \sum_{n=1}^{\infty} 2^{-n} \varphi_n$  is strictly positive.

If  $(X, C, u)$  is an order unit space and  $x, y \in C^\circ$ , then it follows from Lemma 2.1 that there exist  $0 < \alpha \leq \beta$  such that  $\alpha y \leq_C x \leq_C \beta y$ , and hence we can define

$$M(x/y) := \inf\{\beta > 0 : x \leq_C \beta y\} < \infty.$$

On  $C^\circ$  (Birkhoff's version of) Hilbert's metric is given by

$$d_H(x, y) := \log(M(x/y)M(y/x)).$$

It is easy to verify that  $d_H(\lambda x, \mu y) = d_H(x, y)$  for all  $\lambda, \mu > 0$ , and hence  $d_H$  is not a metric. However, it can be shown that  $d_H$  is a metric between pairs of rays in  $C^\circ$ , see [12, Lemma 2.1]. We shall denote the projective spaces obtained by identifying points on rays inside  $C^\circ$  by  $P(C^\circ)$ . So,  $(P(C^\circ), d_H)$  is a metric space.

In case there exists a strictly positive state  $\varphi : X \rightarrow \mathbb{R}$ , we can identify  $P(C^\circ)$  with the cross section

$$\Sigma_\varphi := \{x \in C^\circ : \varphi(x) = 1\}.$$

We shall use the following notation. For  $x \in C^\circ$ , we write  $[x] := x/\varphi(x) \in \Sigma_\varphi$ . The boundary of  $\Sigma_\varphi$  relative to the affine space  $\{x \in X : \varphi(x) = 1\}$  in  $(X, \|\cdot\|_u)$  is denoted by  $\partial\Sigma_\varphi$ . Its closure will be denoted by  $\overline{\Sigma}_\varphi$ .

Let us now prove Theorem 1.1.

*Proof of Theorem 1.1.* Let  $\Omega \subset Y$  be a convex set on which  $\delta_H$  is well defined, i.e.,  $\Omega \cap \ell_{xy}$  is open and bounded for every  $x, y \in \Omega$  with  $x \neq y$ . By translating we may assume that  $0 \in \Omega$  and, by restricting to the span of  $\Omega$ , we may also assume that  $\Omega$  spans  $Y$ .

We claim that these assumptions imply that  $\Omega \cap \ell_{0y}$  is open and bounded for every nonzero  $y \in Y$ . Indeed, let  $y \in Y$  be nonzero, then  $y = \sum_{i=1}^n \lambda_i y_i$  for some  $y_i \in \Omega$  and  $\lambda_i \in \mathbb{R}$ , since  $\Omega$  spans  $Y$ . In this representation we may assume that all  $\lambda_i > 0$ , for if  $\lambda_i < 0$  for some  $i$ , then we can replace  $y_i$  by  $-\varepsilon y_i$  for some small  $\varepsilon$  (since  $\ell_{0y_i} \cap \Omega$  is open) and  $\lambda_i$  by  $-\varepsilon^{-1} \lambda_i > 0$ . Now if  $\lambda := \sum_{i=1}^n \lambda_i$ , then the  $\lambda^{-1} \lambda_i$  are positive and sum to 1, and so

$$\lambda^{-1} y = \sum_{i=1}^n \lambda^{-1} \lambda_i y_i \in \Omega$$

by convexity of  $\Omega$ . So,  $\Omega \cap \ell_{0y} = \Omega \cap \ell_{0(\lambda^{-1}y)}$  is open and bounded.

By considering lines through 0, this implies that  $\Omega$  is absorbing, i.e., for every  $y \in Y$ , there exists an  $\varepsilon > 0$  such that  $\lambda y \in \Omega$  for all  $|\lambda| < \varepsilon$ . Define

$$\overline{\Omega} := \{y \in Y : \lambda y \in \Omega \text{ for all } 0 \leq \lambda < 1\}$$

and note that  $\overline{\Omega}$  is convex.

Let  $X := Y \oplus \mathbb{R}$  and consider the cone

$$C_\Omega := \{\lambda(y, 1) \in X : \lambda \geq 0 \text{ and } y \in \overline{\Omega}\}.$$

We now show that  $C_\Omega$  is Archimedean and  $(0, 1) \in X$  is an order unit.

Let  $(y, \lambda) \in X$ . To show that  $(0, 1) \in X$  is an order unit, we have to find  $M > 0$  with  $-M(0, 1) \leq_{C_\Omega} (y, \lambda) \leq_{C_\Omega} M(0, 1)$ . These inequalities

hold if and only if  $(y, \lambda + M), (-y, M - \lambda) \in C_\Omega$ , which is equivalent to  $y/(M + \lambda), -y/(\lambda - M) \in \overline{\Omega}$ . But the existence of such an  $M$  now follows from the fact that  $\overline{\Omega}$  is absorbing.

To show that  $C_\Omega$  is Archimedean, it suffices to prove that  $n(y, \lambda) \leq_{C_\Omega} (0, 1)$  for all large enough  $n$  implies that  $-(y, \lambda) \in C_\Omega$ , as  $(0, 1)$  is an order unit. The assumed inequality is equivalent with  $(-ny, 1 - n\lambda) \in C_\Omega$  which shows that  $\lambda \leq 0$ . If  $\lambda = 0$ , then  $-ny \in \overline{\Omega}$ , which implies by the boundedness of  $\ell_{0y} \cap \Omega$  that  $y = 0$  and so  $-(y, \lambda) = (0, 0) \in C_\Omega$ . If  $\lambda < 0$ , then by scaling we may assume that  $\lambda = -1$ , and then dividing the inequality by  $1 + n$  yields

$$\frac{n}{1+n}(-y) \in \overline{\Omega}.$$

Hence  $-y \in \overline{\Omega}$ , and so  $-(y, \lambda) = (-y, 1) \in C_\Omega$ .

Obviously, the linear functional  $\varphi: X \rightarrow \mathbb{R}$  defined by  $\varphi((y, s)) := s$  is strictly positive with respect to  $C_\Omega$  and  $\varphi((0, 1)) = 1$ , so  $\varphi$  is a strictly positive state. Moreover,  $\Omega$  is affinely isomorphic to  $\Sigma_\varphi := \{x \in C_\Omega^\circ: \varphi(x) = 1\}$ , which proves the first statement in the theorem.

To prove the second part, we note that for distinct  $x, y \in \Sigma_\varphi$  the points  $w_x := x - M(y/x)^{-1}y$  and  $w_y := y - M(x/y)^{-1}x$  are in  $\partial C \setminus \{0\}$ , as  $C$  is closed. So,  $[w_x]$  and  $[w_y]$  are the end-points of the straight line segment  $\ell_{xy} \cap \Sigma_\varphi$ .  $\square$

*Remark 2.4.* In the proof of Theorem 1.1, the vector space  $Y$  is a subspace of  $X$ , and so it inherits the norm  $\|\cdot\|_u$  from  $X$ . Thus, for  $y \in Y$ ,

$$\|y\| = \inf\{\lambda > 0: -\lambda(0, 1) \leq (y, 0) \leq \lambda(0, 1)\}.$$

The condition on  $\lambda$  is equivalent with  $(y, \lambda), (-y, \lambda) \in C_\Omega$ , which in turn is equivalent with  $y/\lambda, -y/\lambda \in \overline{\Omega}$ , and so

$$\|y\| = \inf\{\lambda > 0: y \in \lambda(\overline{\Omega} \cap -\overline{\Omega})\}.$$

Hence this norm equals the Minkowski functional of the absolutely convex and absorbing set  $\overline{\Omega} \cap -\overline{\Omega}$ .

The advantage of working with cones is that we can use  $d_H$  instead of  $\delta_H$  and apply ideas from the theory of partially ordered vector spaces. Indeed, the following result, which goes back to Birkhoff [2], is well known, see for example [11, Theorem 2.12 and Corollary 2.5.6].

**Lemma 2.5.** *If  $(X, C, u)$  is an order unit space with strictly positive state  $\varphi$ , then on  $\Sigma_\varphi$  the metrics  $d_H$  and  $\delta_H$  coincide. Moreover, the Hilbert's metric topology on  $\Sigma_\varphi$  coincides with the order unit norm topology on  $X$ .*

If  $(X_1, C_1, u_1)$  and  $(X_2, C_2, u_2)$  are two order unit spaces and  $T: X_1 \rightarrow X_2$  is a linear map, we say that  $T$  is *bi-positive* if  $Tx \in C_2$  if and only if  $x \in C_1$ . Note that a bi-positive linear map  $T: X_1 \rightarrow X_2$  is always injective, as  $Tx = 0$  implies that  $x$  and  $-x$  in  $C_1$ , so that  $x = 0$ . Also note that if there exists  $x \in C_1^\circ$  such that  $Tx \in C_2^\circ$ , then  $T(C_1^\circ) \subseteq C_2^\circ$ . Indeed, if  $y \in C_1^\circ$ , then  $y$  is an

order unit by Lemma 2.1, and hence there exists  $\lambda > 0$  such that  $x \leq_{C_1} \lambda y$ . It follows that  $Tx \leq_{C_2} \lambda Ty$ , so that  $Ty$  is an order unit, as  $Tx$  is an order unit, and hence  $Ty \in C_2^\circ$ .

It is easy to check that a bi-positive linear map  $T: X_1 \rightarrow X_2$  induces a Hilbert's metric isometry, as  $M(Tx/Ty) = M(x/y)$  for all  $x, y \in C_1^\circ$ . In that case we shall denote the induced map between the projective spaces  $P(C_1^\circ)$  and  $P(C_2^\circ)$  by  $[T]$ . An isometry  $f$  from  $(P(C_1^\circ), d_H)$  into  $(P(C_2^\circ), d_H)$  is called a *projective linear homomorphism* if there exists a bi-positive linear map  $T: X_1 \rightarrow X_2$  such that  $f = [T]$  on  $P(C_1^\circ)$ .

Before we start the proof of Theorem 1.2, we collect some final pieces of notation. Recall that the image of map  $\gamma$  from a (possibly unbounded) interval  $I \subseteq \mathbb{R}$  into  $(\Sigma_\varphi, d_H)$  is a *geodesic* if

$$d_H(\gamma(t), \gamma(s)) = |t - s| \text{ for all } t, s \in I.$$

Given a straight line  $\ell_{xy}$  through  $x \neq y$  in  $\Sigma_\varphi$ , we write  $\ell_{xy}^+ := \ell_{x,y} \cap \Sigma_\varphi$ . Also for  $x, y \in X$  the closed and open line segments are, respectively, denoted by

$$[x, y] := \{tx + (1-t)y : 0 \leq t \leq 1\} \quad \text{and} \quad (x, y) := \{tx + (1-t)y : 0 < t < 1\}.$$

The half-open intervals are defined in a similar way. For each  $x, y \in \Sigma_\varphi$  the segment  $[x, y]$  is a geodesic in  $(\Sigma_\varphi, d_H)$ , see [9]. It is, however, in general not the only geodesic. The unique geodesics in  $(\Sigma_\varphi, d_H)$  are characterised as follows, see [9] or [10, Proposition 2].

**Lemma 2.6.** *Let  $(X, C, u)$  be an order unit space with strictly positive functional  $\varphi$ . If  $x, y \in \Sigma_\varphi$  and  $x', y' \in \partial\Sigma_\varphi$  are the end points of  $\ell_{xy} \cap \Sigma_\varphi$ , then  $[x, y]$  is the unique geodesic connecting  $x$  and  $y$  in  $(\Sigma_\varphi, d_H)$  if and only if there exist no open line segments  $I_{x'}$  through  $x'$  and  $I_{y'}$  through  $y'$  in  $\partial\Sigma_\varphi$  such that the affine span of  $I_{x'} \cup I_{y'}$  is 2-dimensional.*

### 3. STRICTLY CONVEX HILBERT GEOMETRIES

In this section we prove Theorem 1.2. Throughout this section we shall assume that  $(X_1, C_1, u_1)$  and  $(X_2, C_2, u_2)$  are order unit spaces, with strictly positive states  $\varphi_i: X_i \rightarrow \mathbb{R}$  for  $i = 1, 2$ . For simplicity we write

$$\Sigma_i := \{x \in C_i^\circ : \varphi_i(x) = 1\} \quad \text{and} \quad \|\cdot\|_i := \|\cdot\|_{u_i}$$

for  $i = 1, 2$ .

Recall that  $\Sigma_i$ ,  $i = 1, 2$ , is *strictly convex* if for each  $x, y \in \partial\Sigma_i$  we have that  $(x, y) \subseteq \Sigma_i$ . In that case, it follows from Lemma 2.6 that the metric spaces  $(\Sigma_i, d_H)$ ,  $i = 1, 2$ , are uniquely geodesic. Thus, any isometry  $f$  of  $(\Sigma_1, d_H)$  into  $(\Sigma_2, d_H)$  has to map line segments to line segments if  $(\Sigma_1, d_H)$  and  $(\Sigma_2, d_H)$  are strictly convex. We shall show that any isometry that maps line segments to line segments must be a projective linear homomorphism, which implies Theorem 1.2. We begin with the following lemma.

**Lemma 3.1.** *If  $f: (\Sigma_1, d_H) \rightarrow (\Sigma_2, d_H)$  is an isometry that maps line segments to line segments, then  $f$  has a unique continuous extension to a map from  $(\overline{\Sigma}_1, \|\cdot\|_1)$  into  $(\overline{\Sigma}_2, \|\cdot\|_2)$ . Furthermore, this extension is injective.*

*Proof.* The uniqueness of the continuous extension follows from the density of  $\Sigma_1$  in  $\overline{\Sigma}_1$ . Convergence in this proof will always be in the order unit norm; recall that on  $\Sigma_i$ , the norm and  $d_H$ -topology coincide by Lemma 2.5.

Fix an element  $p \in \Sigma_1$  and let  $x \in \partial\Sigma_1$ . The line segment  $[p, x]$  is mapped onto the line segment  $[f(p), \xi]$ , for some  $\xi \in \partial\Sigma_2$ , because

$$d_H(p, (1-t)p + tx) \rightarrow \infty$$

as  $t \uparrow 1$ . We define  $f(x) := \xi$ .

To show continuity of the extension, let  $(x_n)_n$  be a sequence in  $\overline{\Sigma}_1$  converging to  $x$ . Then  $y_n := (p + x_n)/2 \rightarrow (p + x)/2 =: y$ , and so  $f(y_n) \rightarrow f(y)$ . Let  $s > 1$  be such that  $sf(y) + (1-s)f(p) \notin \overline{\Sigma}_2$ . Then for some  $N \geq 1$  and all  $n \geq N$ ,

$$(3.1) \quad sf(y_n) + (1-s)f(p) \notin \overline{\Sigma}_2.$$

Since  $f$  maps line segments to line segments, there exist  $s_n$  such that

$$s_nf(y_n) + (1-s_n)f(p) = f(x_n) \in \overline{\Sigma}_2,$$

and combining this with (3.1) yields  $s_n < s$  for all  $n \geq N$ .

Now suppose that  $f(x_n)$  does not converge to  $f(x)$ . By passing to a subsequence three times, we find a subsequence  $(x_{n_k})_k$  such that  $f(x_{n_k})$  stays away from  $f(x)$ , that  $s_{n_k} \rightarrow r \in [0, s]$ , and that either all  $x_{n_k} \in \partial\Sigma_1$  or all  $x_{n_k} \in \Sigma_1$ . It follows that

$$f(x_{n_k}) = s_{n_k}f(y_{n_k}) + (1-s_{n_k})f(p) \rightarrow rf(y) + (1-r)f(p) \in \overline{\Sigma}_2.$$

We claim that  $rf(y) + (1-r)f(p) \in \partial\Sigma_2$ . Indeed, if all  $x_{n_k} \in \partial\Sigma_1$ , then  $f(x_{n_k}) \in \partial\Sigma_2$  and the claim follows from the closedness of  $\partial\Sigma_2$ . If all  $x_{n_k} \in \Sigma_1$ , then the fact  $f$  is a  $d_H$ -isometry combined with  $d_H(x_{n_k}, p) \rightarrow \infty$  yields  $d_H(f(x_{n_k}), f(p)) \rightarrow \infty$ , and so  $rf(y) + (1-r)f(p) \in \partial\Sigma_2$ .

Hence  $f(x) = rf(y) + (1-r)f(p)$  by construction of  $f(x)$ , and so  $f(x_{n_k}) \rightarrow f(x)$ , which is impossible since  $f(x_{n_k})$  stays away from  $f(x)$ . Therefore the extension of  $f$  is continuous.

To show injectivity suppose that  $x, y \in \overline{\Sigma}_1$  are such that  $x \neq y$ . Note that  $[f(p), f(x)] \neq [f(p), f(y)]$ , since  $f$  is injective on  $\Sigma_1$ , so that  $f(x) \neq f(y)$ .  $\square$

The following lemma is elementary, but useful in the sequel.

**Lemma 3.2.** *Let  $x, y \in \Sigma_1$  and let  $x', y' \in \partial\Sigma_1$  be the end points of  $\ell_{xy}^+$  such that  $x$  is between  $x'$  and  $y$ , and  $y$  is between  $y'$  and  $x$ . Suppose that  $f: (\Sigma_1, d_H) \rightarrow (\Sigma_2, d_H)$  is an isometry that maps line segments to line segments. If  $f(x)'$  and  $f(y)'$  are the end points of  $\ell_{f(x)f(y)}^+$  such that  $f(x)$  is between  $f(x)'$  and  $f(y)$ , and  $f(y)$  is between  $f(y)'$  and  $f(x)$ , then there exists a linear map  $S: \text{Span}\{x', y'\} \rightarrow \text{Span}\{f(x)', f(y)'\}$  satisfying  $[Sx] = f(x)'$ ,*



$[Sy'] = f(y)'$ ,  $[Sz] = f(z)$  for all  $z \in (x', y')$ , and  $S$  is bi-positive with respect to  $C_1 \cap \text{Span}\{x', y'\}$  and  $C_2 \cap \text{Span}\{f(x)', f(y)'\}$ .

*Proof.* Let  $0 < s, t < 1$  be such that  $x = tx' + (1-t)y'$  and  $f(x) = sf(x') + (1-s)f(y')$ . Since  $\{tx', (1-t)y'\}$  and  $\{sf(x'), (1-s)f(y)'\}$  define bases for  $\text{Span}\{x', y'\}$  and  $\text{Span}\{f(x)', f(y)'\}$ , and the cones  $C_1 \cap \text{Span}\{x', y'\}$  and  $C_2 \cap \text{Span}\{f(x)', f(y)'\}$  are the positive span of these basis elements, we have a bijective and bi-positive linear map  $S: \text{Span}\{x', y'\} \rightarrow \text{Span}\{f(x)', f(y)'\}$  defined by

$$S(\alpha tx' + \beta(1-t)y') := \alpha sf(x') + \beta(1-s)f(y').$$

Note that  $[Sx'] = f(x)'$ ,  $[Sy'] = f(y)'$  and  $Sx = f(x)$ .

Let  $z \in [x, x')$  and note that, as  $f$  maps line segments to line segments and  $f$  is an isometry, we have that

$$[f(x)', f(z), f(x), f(y)'] = \exp(d_H(f(z), f(x))) = \exp(d_H(z, x)) = [x', z, x, y'].$$

As the cross ratio is a projective invariant, we know that

$$[[Sx'], [Sz], [Sx], [Sy']] = [Sx', Sx, Sy, Sy'] = [x', z, x, y'],$$

which combined with the previous equality gives

$$[[Sx'], [Sz], [Sx], [Sy']] = [f(x)', f(z), f(x), f(y)'] = [[Sx'], f(z), [Sx], [Sy']] .$$

This implies that  $[Sz] = f(z)$ . Interchanging the roles of  $x'$  and  $y'$  finally gives  $[Sz] = f(z)$  for all  $z \in (x', y')$ .  $\square$

**Proposition 3.3.** *Let  $f: (\Sigma_1, d_H) \rightarrow (\Sigma_2, d_H)$  be an isometry that maps line segments to line segments and  $Y \subseteq X_1$  be a subspace such that  $Y \cap \Sigma_1 \neq \emptyset$ . If  $T: Y \rightarrow X_2$  is a bi-positive linear map such that  $[Ty] = f(y)$  for all  $y \in Y \cap \Sigma_1$ , then for  $z \in \Sigma_1 \setminus Y$  there is a bi-positive linear extension*

$$\hat{T}: Y \oplus \text{Span}\{z\} \rightarrow X_2$$

*of  $T$  such that  $[\hat{T}y] = f(y)$  for all  $y \in (Y \oplus \text{Span}\{z\}) \cap \Sigma_1$ .*

*Proof.* Let  $T: Y \rightarrow X_2$  be a bi-positive linear map that satisfies  $[Ty] = f(y)$  for all  $y \in Y \cap \Sigma_1$  and choose  $z \in \Sigma_1 \setminus Y$ . Fix  $\xi \in Y \cap \Sigma_1$  and consider  $\ell_{z, \xi}^+$ . By Lemma 3.2 there exists a bi-positive linear map  $S: \text{Span}\{\xi, z\} \rightarrow X_2$  such that  $[Sx] = f(x)$  for all  $x \in \ell_{z, \xi}^+$ . By rescaling  $S$ , we may assume that  $S\xi = T\xi$ .

Now let  $\hat{Y} := Y \oplus \text{Span}\{z\}$  and define the linear map

$$\hat{T}: \hat{Y} \rightarrow \text{ran } T + \text{Span}\{Sz\}$$

by  $y + \lambda z \mapsto Ty + \lambda Sz$ . We wish to show that  $\text{ran } T + \text{Span}\{Sz\}$  is in fact a direct sum, making  $\hat{T}$  injective. Suppose  $Ty = \lambda Sz$  for some  $y \in Y$  and  $\lambda > 0$ . As  $T$  is bi-positive, this implies that  $y \in Y \cap C_1^\circ$ , so  $[Ty] = [Sz] = f(z)$ . But this yields  $f([y]) = f(z)$ , so  $[y] = z$  and this is impossible. If  $\lambda < 0$ , we can use a similar argument for  $-y$  to arrive at a contradiction.

Thus,  $\text{ran } T + \text{Span}\{Sz\}$  is indeed a direct sum. Since  $\hat{T}\xi = T\xi = S\xi$  and  $\hat{T}z = Sz$ , it follows that  $\hat{T} = S$  on  $\ell_{z,\xi}^+$  and so  $[\hat{T}x] = f(x)$  for all  $x \in \ell_{z,\xi}^+$ .

Now suppose that  $w \in \hat{Y} \cap \Sigma_1$  and  $w \notin Y \cup \ell_{z,\xi}^+$ . The subspace  $Y$  is a hyperplane in  $\hat{Y}$  and therefore, it divides  $\hat{Y} \cap \Sigma_1$  into two parts. Choose distinct  $\eta_1, \eta_2 \in \ell_{z,\xi}^+$  that lie on the other side of  $w$  and let  $y_1, y_2 \in Y \cap \Sigma_1$  be the intersection points of the line segments  $\ell_{w,\eta_1}^+$  and  $\ell_{w,\eta_2}^+$ , respectively. The situation is depicted in Figure 1 below. We see that  $w$  is the unique

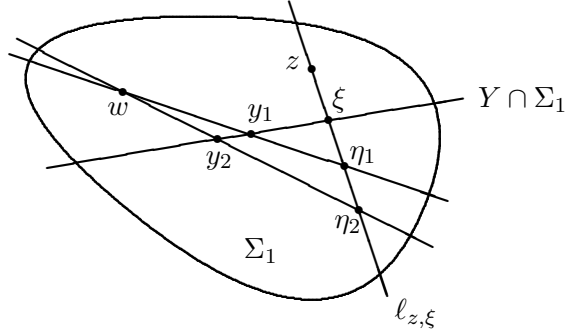


FIGURE 1. Points of intersection

point of intersection of the line segments  $\ell_{\eta_1,y_1}^+$  and  $\ell_{\eta_2,y_2}^+$ . As  $f$  is injective, it follows that  $f(w)$  is the unique point of intersection of the line segments  $\ell_{f(\eta_1),f(y_1)}^+$  and  $\ell_{f(\eta_2),f(y_2)}^+$ . Also, as  $[\hat{T}\eta_i] = f(\eta_i)$  and  $[\hat{T}y_i] = f(y_i)$ , we have  $\hat{T}\eta_i, \hat{T}y_i \in \text{Span}\{f(\eta_i), f(y_i)\}$ ; hence

$$\hat{T}w \in \text{Span}\{f(\eta_1), f(y_1)\} \cap \text{Span}\{f(\eta_2), f(y_2)\} = \text{Span}\{f(w)\}.$$

So,  $\hat{T}w \in C_2^\circ \cup -C_2^\circ$ , since  $\hat{T}$  is injective, and hence

$$\hat{T}(\hat{Y} \cap C_1^\circ) \subseteq C_2^\circ \cup -C_2^\circ.$$

The convexity of  $\hat{Y} \cap C_1^\circ$  now implies that either  $\hat{T}(\hat{Y} \cap C_1^\circ) \subseteq C_2^\circ$  or  $\hat{T}(\hat{Y} \cap C_1^\circ) \subseteq -C_2^\circ$ . As  $\hat{T}\xi = T\xi \in C_2^\circ$ , it follows that  $\hat{T}(\hat{Y} \cap C_1^\circ) \subseteq C_2^\circ$ . Moreover, this shows that  $[\hat{T}w]$  is well defined. As  $w$  was arbitrary, we conclude that  $[\hat{T}y] = f(y)$  for all  $y \in \hat{Y} \cap \Sigma_1$ .

Next, we will show that  $\hat{T}$  is bi-positive. Suppose  $x \in \partial C_1 \cap \hat{Y}$  with  $x \neq 0$ . We claim that  $\hat{T}(x) \in \partial C_2$ . Since  $\varphi_1(x) > 0$ , we may assume without loss of generality that  $x \in \partial \Sigma_1$ . Define  $x_n := (1 - \frac{1}{n})x + \frac{1}{n}\xi$  for  $n \geq 1$ . Then we have  $x_n \rightarrow x$ , so that  $f(x_n) \rightarrow f(x)$  by Lemma 3.1 and

$$\hat{T}x_n = \hat{T}|_{\text{Span}\{x,\xi\}}x_n \rightarrow \hat{T}|_{\text{Span}\{x,\xi\}}x = \hat{T}x$$

as  $\dim(\text{Span}\{x,\xi\}) < \infty$ . It follows that  $\varphi_2(\hat{T}x_n) \rightarrow \varphi_2(\hat{T}x)$  and, since  $\hat{T}x_n \in C_2$  for all  $n \geq 1$ , we must have  $\hat{T}x \in C_2$ , because  $C_2$  is closed. Moreover, the injectivity of  $\hat{T}$  yields  $\varphi_2(\hat{T}x) > 0$ . We also have that

$$\hat{T}x_n = \varphi_2(\hat{T}x_n)f(x_n) \rightarrow \varphi_2(\hat{T}x)f(x) \in \partial C_2,$$

so we conclude that  $\hat{T}x \in \partial C_2$  showing that  $\hat{T}$  is positive. Finally, if we pick  $x \in \hat{Y} \setminus C_1$ , then there exists a  $0 < t < 1$  such that  $tx + (1-t)\xi \in \partial C_1$  and so  $t\hat{T}x + (1-t)\hat{T}\xi \in \partial C_2$  by our previous findings. But this implies that  $\hat{T}x \notin C_2$  making  $\hat{T}$  bi-positive.  $\square$

Enough preparations have been made to prove the following result which implies Theorem 1.2 by Theorem 1.1.

**Theorem 3.4.** *An map  $f: (\Sigma_1, d_H) \rightarrow (\Sigma_2, d_H)$  is an isometry that maps line segments to line segments if and only if it is a projective linear homomorphism.*

*Proof.* Consider the collection  $\mathcal{C}$  of pairs  $(Y, T_Y)$  where  $Y \subseteq X_1$  is a linear subspace and  $T_Y: Y \rightarrow X_2$  is a bi-positive linear map such that  $[T_Y x] = f(x)$  for all  $x \in Y \cap \Sigma_1$ . Note that  $\mathcal{C} \neq \emptyset$ , since  $C_1^\circ \neq \emptyset$ . We can define a partial order  $\leq$  on  $\mathcal{C}$  by

$$(Y, T_Y) \leq (Z, T_Z) \text{ if } Y \subseteq Z \text{ and } T_Z y = T_Y y \text{ for all } y \in Y.$$

Let  $(Y_i, T_{Y_i})_{i \in I}$  be a totally ordered subset in  $\mathcal{C}$ . Put  $Y := \bigcup_{i \in I} Y_i$  and define  $T_Y: Y \rightarrow X_2$  through  $T_Y y_i := T_{Y_i} y_i$ . Clearly, we have that  $Y$  is a linear subspace of  $X_1$  and  $T_Y$  is a well defined bi-positive linear map. For  $y \in Y \cap \Sigma_1$  we have  $y \in Y_i \cap \Sigma_1$  for some  $i \in I$  and  $[T_Y y] = [T_{Y_i} y] = f(y)$ , so  $(Y, T_Y) \in \mathcal{C}$  is an upper bound. By Zorn's lemma our collection  $\mathcal{C}$  contains a maximal element  $(\Omega, T_\Omega)$ . Suppose that  $x \in \Sigma_1 \setminus \Omega$ . By Proposition 3.3 we have a bi-positive linear extension

$$\hat{T}_\Omega: \Omega \oplus \text{Span}\{x\} \rightarrow X_2$$

of  $T_\Omega$  such that  $[\hat{T}_\Omega z] = f(z)$  for all  $z \in (\Omega \oplus \text{Span}\{x\}) \cap \Sigma_1$ . But now we have  $(\Omega \oplus \text{Span}\{x\}, \hat{T}_\Omega) \in \mathcal{C}$  and  $(\Omega, T_\Omega) \leq (\Omega \oplus \text{Span}\{x\}, \hat{T}_\Omega)$ , which contradicts the maximality of  $(\Omega, T_\Omega)$ . We conclude that  $\Omega \cap \Sigma_1 = \Sigma_1$  and therefore  $[T_\Omega x] = f(x)$  for all  $x \in \Sigma_1$ .

We claim that  $\Omega = X_1$ . Let  $x \in X_1$ ,  $u \in \Sigma_1$  and  $\varepsilon > 0$  such that  $B_\varepsilon(u) \subseteq C_1^\circ$ . Also, there are  $y, z \in C_1$  such that  $x = y - z$ , as  $C_1 - C_1 = X_1$ . Now, for  $k \geq 1$  large enough, we have that  $\frac{1}{k}y + u, \frac{1}{k}z + u \in B_\varepsilon(u)$ . Since  $\Sigma_1 \subseteq \Omega$ , it follows that  $C_1^\circ \subseteq \Omega$ ; hence

$$\frac{1}{k}x = \left(\frac{1}{k}y + u\right) - \left(\frac{1}{k}z + u\right) \in \Omega,$$

so  $x \in \Omega$ .

The converse implication is well known.  $\square$

Note that Lemma 2.2 shows that the linear map  $T$  in the previous theorem is continuous with respect to the order unit norm. We also like to note that Theorem 3.4 implies that two uniquely geodesic Hilbert geometries  $\Omega_1$  and  $\Omega_2$  are isometric if and only if there exists a projective linear isomorphism between them.

*Remark 3.5.* An important variant of Hilbert's metric is Thompson's metric which was introduced in [18]. On the interior of a closed cone  $C$ , *Thompson's metric* is given by

$$d_T(x, y) := \log \max\{M(x/y), M(y/x)\} \quad \text{for } x, y \in C^\circ.$$

It was shown in [13, Theorem 8.2] that if  $C$  is a finite dimensional strictly convex cone with  $\dim C \geq 3$ , then for every isometry  $f$  of  $(C^\circ, d_T)$  there exists a bi-positive linear map  $T: X \rightarrow X$  such that for each  $x \in C^\circ$  we have that  $f(x) = \lambda_x T x$  for some  $\lambda_x > 0$ . The proof of this result relies on [10, Proposition 3]. Using Theorem 3.4 it is straightforward to extend [13, Theorem 8.2] to infinite dimensional strictly convex cones.

#### 4. INFINITE DIMENSIONAL SIMPLICES

Let  $K$  be a compact Hausdorff space and  $C(K)$  denote the space of real-valued continuous functions on  $K$ . Consider the cone  $C(K)_+$  consisting of nonnegative functions with interior,

$$C(K)_+^\circ := \{f \in C(K) : f(x) > 0 \text{ for all } x \in K\}.$$

It is well known that the Hilbert geometry on a finite dimensional simplex is isometric to a finite dimensional normed space. The same is true for  $(P(C(K)_+^\circ), d_H)$ , see [17, Proposition 1.7]. It will be useful to recall the basic argument. Let  $\mathbf{1}$  be the constant one function on  $K$  and denote the elements in the quotient space  $\overline{C(K)} := C(K)/\mathbb{R}\mathbf{1}$  by  $\bar{g}$ . The map  $\text{Log}: P(C(K)_+^\circ) \rightarrow \overline{C(K)}$  given by,

$$\text{Log}(\bar{f}) := \overline{\log \circ f} \quad \text{for } \bar{f} \in P(C(K)_+^\circ),$$

is an isometry of  $(P(C(K)_+^\circ), d_H)$  onto  $(\overline{C(K)}, \|\cdot\|_{\text{var}})$ , where

$$\|\bar{g}\|_{\text{var}} := \sup_{x \in K} g(x) - \inf_{x \in K} g(x),$$

is the *variation norm*. To see this note that

$$M(f/g) = \inf\{\beta > 0 : f(x) \leq \beta g(x) \text{ for all } x \in K\} = \sup_{x \in K} \frac{f(x)}{g(x)}$$

for  $f, g \in P(C(K)_+^\circ)$ , so that

$$\begin{aligned} d_H(f, g) &= \sup_{x \in K} \frac{\log f(x)}{\log g(x)} + \sup_{x \in K} \frac{\log g(x)}{\log f(x)} \\ &= \sup_{x \in K} (\log f(x) - \log g(x)) - \inf_{x \in K} (\log f(x) - \log g(x)) \\ &= \|\text{Log}(\bar{f}) - \text{Log}(\bar{g})\|_{\text{var}}. \end{aligned}$$

Given a function  $f \in C(K)$  the supremum norm of its translations  $f - \lambda \mathbf{1}$  for  $\lambda \in \mathbb{R}$  is minimised precisely when translating  $f$  by the average of both extreme values  $\frac{1}{2}(\sup_{x \in K} f(x) + \inf_{x \in K} f(x))$ . This means that the quotient norm on  $C(K)/\mathbb{R}\mathbf{1}$  with respect to the supremum norm is exactly half of  $\|\cdot\|_{\text{var}}$ . This assertion is made precise in the following lemma.

**Lemma 4.1.** *Let  $K$  be a compact Hausdorff space. If  $\|\cdot\|_q$  is the quotient norm on  $\overline{C(K)}$  with respect to  $2\|\cdot\|_\infty$ , then  $\|\cdot\|_q$  coincides with  $\|\cdot\|_{\text{var}}$  on  $\overline{C(K)}$ .*

*Proof.* Let  $f \in C(K)$  and write  $a \vee b = \max\{a, b\}$  for  $a, b \in \mathbb{R}$ . Using the elementary fact that  $\inf_{\lambda \in \mathbb{R}} (a - \lambda) \vee (b + \lambda) = (a + b)/2$ , we see that

$$\begin{aligned} \|\overline{f}\|_q &= 2 \inf_{\lambda \in \mathbb{R}} \|f - \lambda \mathbf{1}\|_\infty = 2 \inf_{\lambda \in \mathbb{R}} \sup_{s \in K} |f(s) - \lambda| \\ &= 2 \inf_{\lambda \in \mathbb{R}} \left[ \left( \sup_{s \in K} f(s) - \lambda \right) \vee \left( \sup_{s \in K} -f(s) + \lambda \right) \right] \\ &= \sup_{s \in K} f(s) + \sup_{s \in K} -f(s) = \sup_{s \in K} f(s) - \inf_{s \in K} f(s) = \|\overline{f}\|_{\text{var}}. \end{aligned}$$

□

Our findings so far have shown that in order to describe the surjective isometries

$$h: (P(C(K_1)_+^\circ), d_H) \rightarrow (P(C(K_2)_+^\circ), d_H),$$

it suffices to understand the surjective isometries

$$T: (\overline{C(K_1)}, \|\cdot\|_q) \rightarrow (\overline{C(K_2)}, \|\cdot\|_q).$$

By the Mazur-Ulam theorem these isometries  $T$  must be affine. We can compose  $T$  with an appropriate translation to make it linear. Thus, our goal will be to classify all isometric isomorphisms (surjective linear isometries)  $T: (\overline{C(K_1)}, \|\cdot\|_q) \rightarrow (\overline{C(K_2)}, \|\cdot\|_q)$ .

We will follow the lines of the proof for the Banach-Stone theorem [5, Theorem VI.2.1]. The Banach-Stone theorem characterises the isometric isomorphisms between  $(C(K), \|\cdot\|_\infty)$  spaces. A common way to prove the Banach-Stone theorem is by looking at the adjoint operator, which is an isometry on the dual of  $(C(K), \|\cdot\|_\infty)$ , and to exploit the extreme points of the unit ball there. We shall take a similar approach.

Recall that the dual of  $(C(K), \|\cdot\|_\infty)$  is  $(M(K), \|\cdot\|_{TV})$ , where  $M(K)$  is the space of all regular Borel measures on  $K$  and  $\|\mu\|_{TV} := |\mu|(K)$  is the *total variation norm*. Let us recall some basic facts about  $(M(K), \|\cdot\|_{TV})$ , which can be found in [5, Appendix C].

Every  $\mu \in M(K)$  has a Hahn-Jordan decomposition  $\mu = \mu^+ - \mu^-$  where  $\mu^+$  and  $\mu^-$  are positive measures in  $M(K)$ , and

$$(4.1) \quad \|\mu\|_{TV} = \|\mu^+\|_{TV} + \|\mu^-\|_{TV}.$$

Also, if  $\mu, \nu \in M(K)$  are positive, we have  $\|\mu + \nu\|_{TV} = \|\mu\|_{TV} + \|\nu\|_{TV}$ . The space  $M(K)$  is a lattice: every  $\mu, \nu \in M(K)$  have a supremum (least upper bound)  $\mu \vee \nu$ . The set  $P(K)$  denotes the set of probability measures on  $K$ , and its extreme points are the set of Dirac measures  $\{\delta_s : s \in K\}$ . The map  $s \mapsto \delta_s$  is a homeomorphism from  $K$  onto  $\{\delta_s : s \in K\}$  equipped with the weak\*-topology.

The dual space of  $(C(K)/\mathbb{R}\mathbf{1}, \|\cdot\|_q)$  is  $\mathbb{R}\mathbf{1}^\perp \subseteq (M(K), \frac{1}{2}\|\cdot\|_{TV})$ , where  $\mathbb{R}\mathbf{1}^\perp := \{\mu \in M(K) : \int_K \alpha \mathbf{1} d\mu = 0 \text{ for all } \alpha \in \mathbb{R}\}$ . It follows that

$$\mathbb{R}\mathbf{1}^\perp = \{\mu \in M(K) : \|\mu^+\|_{TV} = \|\mu^-\|_{TV}\}.$$

Now, if

$$T : (\overline{C(K_1)}, \|\cdot\|_q) \rightarrow (\overline{C(K_2)}, \|\cdot\|_q)$$

is an isometric isomorphism, then the corresponding adjoint operator

$$T^* : (\mathbb{R}\mathbf{1}_2^\perp, \frac{1}{2}\|\cdot\|_{TV}) \rightarrow (\mathbb{R}\mathbf{1}_1^\perp, \frac{1}{2}\|\cdot\|_{TV})$$

is a isometric isomorphism as well. Moreover,  $T^*$  is a weak\*-homeomorphism from the unit ball  $B_2 \subseteq \mathbb{R}\mathbf{1}_2^\perp$  onto the unit ball  $B_1 \subseteq \mathbb{R}\mathbf{1}_1^\perp$  that maps the set of extreme points,  $\text{ext}(B_2)$ , of  $B_2$  bijectively onto the set of extreme points,  $\text{ext}(B_1)$ , of  $B_1$ . The following lemma tells us that the extreme points are exactly the differences of Dirac measures.

**Proposition 4.2.** *Let  $K$  be a compact Hausdorff space. The set of extreme points,  $\text{ext}(B)$ , of the unit sphere  $B$  in  $\mathbb{R}\mathbf{1}^\perp \subseteq (M(K), \frac{1}{2}\|\cdot\|_{TV})$  satisfies*

$$\text{ext}(B) = \{\delta_s - \delta_t : s, t \in K \text{ and } s \neq t\}.$$

*Proof.* Let  $\delta_s - \delta_t \in B$  with  $s \neq t$  and suppose  $\mu, \nu \in B$  are such that

$$\delta_s - \delta_t = \frac{1}{2}(\mu + \nu) = \frac{1}{2}(\mu^+ + \nu^+) - \frac{1}{2}(\mu^- + \nu^-).$$

Then  $\delta_s = (\delta_s - \delta_t) \vee 0 \leq \frac{1}{2}(\mu^+ + \nu^+)$ , and so for  $\eta := \frac{1}{2}(\mu^+ + \nu^+) - \delta_s \geq 0$  it follows from (4.1) that

$$1 + \|\eta\|_{TV} = \|\delta_s + \eta\|_{TV} = \frac{1}{2}\|\mu^+ + \nu^+\|_{TV} = \frac{1}{2}\|\mu^+\|_{TV} + \frac{1}{2}\|\nu^+\|_{TV} = 1,$$

so  $\eta = 0$ , which yields  $\delta_s = \frac{1}{2}(\mu^+ + \nu^+)$ . The fact that  $\delta_s$  is an extreme point in  $B_{M(K)}$  (see [5, Theorem V.8.4]) implies that  $\mu^+ = \nu^+$ . Similarly, we have that  $\delta_t \leq \frac{1}{2}(\mu^- + \nu^-)$  which implies  $\mu^- = \nu^-$ . This shows that  $\delta_s - \delta_t \in \text{ext}(B)$ .

Conversely, let  $\mu = \mu^+ - \mu^- \in \text{ext}(B)$ , and suppose  $\mu^+ = \frac{1}{2}(\nu_1 + \nu_2)$  for  $\nu_1, \nu_2 \in P(K)$ . Then for  $i = 1, 2$ ,

$$\frac{1}{2}\|\nu_i - \mu^-\|_{TV} \leq \frac{1}{2}(\|\nu_i\|_{TV} + \|\mu^-\|_{TV}) = 1,$$

so  $\nu_i - \mu^- \in B$ . Since  $\mu \in \text{ext}(B)$  and  $\mu = \frac{1}{2}(\nu_1 - \mu^-) + \frac{1}{2}(\nu_2 - \mu^-)$ , we obtain  $\nu_1 = \nu_2$ . Hence  $\mu^+ \in \text{ext}(P(K))$ , which implies that  $\mu^+ = \delta_s$  for some  $s \in K$  by [5, Theorem V.8.4]. A similar argument yields  $\mu^- = \delta_t$  for some  $t \in K$ .  $\square$

For  $s \in K$  define the sets  $E_s := \{\delta_s - \delta_t : t \in K, s \neq t\}$ . Clearly the distance between distinct elements of  $E_s$  is 1, and it turns out that the sets  $\pm E_s$  are the maximal equilateral subsets in  $\text{ext}(B)$  of mutual distance 1.

**Lemma 4.3.** *Let  $K$  be a compact Hausdorff space and  $\emptyset \neq A \subseteq \text{ext}(B)$  be such that  $\frac{1}{2}\|\mu - \nu\|_{TV} = 1$  for all  $\mu, \nu \in A$  with  $\mu \neq \nu$ , then there is an element  $s \in K$  such that  $A \subseteq E_s$  or  $A \subseteq -E_s$ .*

*Proof.* If such an  $s$  does not exist, then there exist elements  $\delta_s - \delta_t, \delta_p - \delta_q \in A$  with  $s \neq p$  and  $t \neq q$ . But clearly

$$\frac{1}{2} \|(\delta_s - \delta_t) - (\delta_p - \delta_q)\|_{TV} = 2.$$

□

Now let  $T: \overline{C(K_1)} \rightarrow \overline{C(K_2)}$  be an isometric isomorphism. Then the adjoint  $T^*$  is an isometric isomorphism as well, and so  $T^*$  preserves the maximal equilateral subsets of the extreme points of mutual distance 1. Hence  $T^*(E_s) = \pm E_{\vartheta(s)}$  for some bijection  $\vartheta: K_2 \rightarrow K_1$ . Furthermore, if  $s \neq t$ , then  $E_s \cap E_t = \emptyset$ . But  $E_s \cap -E_t \neq \emptyset$ , as  $\delta_s - \delta_t \in E_s \cap -E_t$ . As  $T^*$  maps disjoint sets to disjoint sets, either  $T^*(E_s) = E_{\vartheta(s)}$  for all  $s \in K_2$  or  $T^*(E_s) = -E_{\vartheta(s)}$  for all  $s \in K_2$ . So  $T^*(E_s) = \varepsilon E_{\vartheta(s)}$  for some  $\varepsilon \in \{-1, 1\}$ .

**Lemma 4.4.** *The above constructed bijection  $\vartheta: K_2 \rightarrow K_1$  is a homeomorphism.*

*Proof.* Let  $(s_\alpha)_\alpha$  be a net in  $K_2$  converging to  $s \in K_2$ . Then  $\delta_{s_\alpha}$  converges weak\* to  $\delta_s$ , and so

$$\delta_{\vartheta(s_\alpha)} - \delta_{\vartheta(s)} = \varepsilon T^*(\delta_{s_\alpha} - \delta_s) \xrightarrow{\text{weak}^*} \varepsilon T^*0 = 0.$$

Hence  $\delta_{\vartheta(s_\alpha)} \xrightarrow{\text{weak}^*} \delta_{\vartheta(s)}$ , or equivalently,  $\vartheta(s_\alpha) \rightarrow \vartheta(s)$ . So  $\vartheta$  is a continuous bijection from a compact space into a Hausdorff space and so it is a homeomorphism. □

We are now able to prove our main result on isometric isomorphisms  $T: \overline{C(K_1)} \rightarrow \overline{C(K_2)}$ .

**Theorem 4.5.** *If  $K_1$  and  $K_2$  are compact Hausdorff spaces, then a map  $T: \overline{C(K_1)} \rightarrow \overline{C(K_2)}$  is an isometric isomorphism if and only if there exist an  $\varepsilon \in \{-1, 1\}$  and a homeomorphism  $\vartheta: K_2 \rightarrow K_1$  such that  $T\bar{f} = \varepsilon(f \circ \vartheta)$ .*

*Proof.* Suppose  $T$  is an isometric isomorphism, and let  $\vartheta$  and  $\varepsilon$  be such that  $T^*E_s = \varepsilon E_{\vartheta(s)}$ . Although point-evaluation for elements  $\bar{g} \in \overline{C(K_2)}$  is not well defined, the values  $\bar{g}(s) - \bar{g}(t)$  for  $s, t \in K_2$  are well defined, and the computation

$$\begin{aligned} T\bar{f}(s) - T\bar{f}(t) &= (\delta_s - \delta_t)(T\bar{f}) \\ &= [T^*(\delta_s - \delta_t)](\bar{f}) \\ &= [\varepsilon(\delta_{\vartheta(s)} - \delta_{\vartheta(t)})](\bar{f}) \\ &= \varepsilon(\bar{f}(\vartheta(s)) - \bar{f}(\vartheta(t))) \end{aligned}$$

shows that  $T$  is induced from the map  $f \mapsto \varepsilon(f \circ \vartheta)$  from  $C(K_1)$  into  $C(K_2)$ .

Conversely, if  $\varepsilon \in \{-1, 1\}$  and  $\vartheta: K_2 \rightarrow K_1$  is a homeomorphism, then  $f \mapsto \varepsilon(f \circ \vartheta)$  is an isometric isomorphism between  $C(K_1)$  and  $C(K_2)$  that maps  $\mathbb{R}\mathbf{1}_1$  onto  $\mathbb{R}\mathbf{1}_2$ , and hence it induces an isometric isomorphism between the respective quotient spaces  $\overline{C(K_1)}$  and  $\overline{C(K_2)}$ . □

Our next goal is to describe the group of surjective isometries from  $\overline{C(K)}$  to itself. For any real normed space  $X$ , the Mazur-Ulam theorem shows that the group of surjective isometries from  $X$  to  $X$  is a product of the group of isometric isomorphisms and the group of translations. Clearly these groups have trivial intersection, and it is easily verified that a translation by  $x \in X$  conjugated by an isometric isomorphism  $T$  of  $X$  yields a translation by  $Tx$ . This shows that the group of surjective isometries from  $X$  to  $X$  is a semidirect product of these two groups. In the case of  $X = \overline{C(K)}$ , we know the group of isometric isomorphisms by Theorem 4.5, which yields the following description.

**Proposition 4.6.** *If  $K$  be a compact Hausdorff space, then the group of surjective isometries of  $\overline{C(K)}$  is isomorphic to  $\overline{C(K)} \rtimes (C_2 \times \text{Homeo}(K))$  if and only if  $|K| \geq 3$ , where  $C_2$  is the cyclic group of order 2.*

Note that in the above proposition,  $\vartheta \in \text{Homeo}(K)$  acts on  $\overline{f} \in \overline{C(K)}$  as  $\overline{f \circ \vartheta^{-1}}$  (not as  $\overline{f \circ \vartheta}$ ).

*Proof.* If  $K$  has only 2 elements, then multiplication by  $-1$  coincides with the non-trivial homeomorphism, and if  $|K| \leq 1$ , then multiplication by  $-1$  is the identity. In all other cases there is no overlap, and the result follows from the above discussion.  $\square$

Translating Theorem 4.5 and Proposition 4.6 back to  $(P(C(K_i)_+^\circ), d_H)$  through the pointwise exponential (the inverse of the pointwise logarithm) yields the following.

**Theorem 4.7.** *Let  $K_1, K_2$  be compact Hausdorff spaces. A map*

$$h: (P(C(K_1)_+^\circ), d_H) \rightarrow (P(C(K_2)_+^\circ), d_H)$$

*is a surjective isometry if and only if there exist  $g \in C(K_2)_+^\circ$ ,  $\varepsilon \in \{-1, 1\}$ , and a homeomorphism  $\vartheta: K_2 \rightarrow K_1$  such that*

$$h(\overline{f}) = \overline{g \cdot (f \circ \vartheta)^\varepsilon} \quad \text{for all } \overline{f} \in P(C(K_2)_+^\circ).$$

*If  $K_1 = K_2 = K$  and  $|K| \geq 3$ , then the isometry group is given by*

$$\text{Isom}(P(C(K)_+^\circ), d_H) \cong \overline{C(K)} \rtimes (C_2 \times \text{Homeo}(K)),$$

*where  $C_2$  is the cyclic group of order 2.*

If  $\mu_i$  ( $i \in \{1, 2\}$ ) is a strictly positive measure on  $K_i$ , and we identify  $P(C(K_i)_+^\circ)$  with  $\Delta(K_i, \mu_i)$ , we obtain Theorem 1.3.

*Remark 4.8.* In the above theorem, everything is projectively linear except for the inversion. It follows that if  $K$  is a compact Hausdorff space with at least 3 elements, the index of the collineation group in the isometry group  $\text{Isom}(P(C(K)_+^\circ), d_H)$  equals 2.

Theorem 4.7 has the following interesting consequence.



**Corollary 4.9.** *If  $K_1, K_2$  are compact Hausdorff spaces, the metric spaces  $(P(C(K_1)_+^\circ), d_H)$  and  $(P(C(K_2)_+^\circ), d_H)$  are isometric if and only if  $K_1$  and  $K_2$  are homeomorphic.*

It would be interesting to study non-commutative versions of Theorem 4.7. In particular, one could look at Hilbert's metric isometries on the interior of the cone  $A_+$  of positive self-adjoint elements in a unital  $C^*$ -algebra. In view of the characterisation of Thompson's metric isometries on  $A_+^\circ$  by Hatori and Molnár [8], it seems plausible that each Hilbert's metric isometry  $h: P(A_+^\circ) \rightarrow P(A_+^\circ)$  is of the form  $h(\bar{a}) = \overline{h(e)^{1/2} J(a^\varepsilon) h(e)^{1/2}}$ , where  $\varepsilon \in \{-1, 1\}$ ,  $e$  is the unit in  $A$ , and  $J$  is a Jordan $^*$  isomorphism.

## REFERENCES

- [1] C.D. Aliprantis and R. Tourky, *Cones and duality*. Graduate Studies in Mathematics, 84. American Mathematical Society, Providence, RI, 2007.
- [2] G. Birkhoff, Extensions of Jentzsch's theorems. *Trans. Amer. Math. Soc.* **85**, (1957), 219–277.
- [3] A. Bosché, Symmetric cones, the Hilbert and Thompson metrics, [arXiv:1207.3214](#), 2012.
- [4] M. Burger, A. Iozzi and N. Monod, Equivariant embeddings of trees into hyperbolic spaces, *Int. Math. Res. Not.* **22**, (2005), 1331–1369.
- [5] J.B. Conway, *A course in functional analysis*. Second edition. Graduate Texts in Mathematics, 96. Springer-Verlag, New York, 1990.
- [6] S. Francaviglia and A. Martino, The isometry group of outer space. *Adv. Math.* **231**(3–4), (2012), 1940–1973.
- [7] *Handbook of Hilbert geometry*. Edited by G. Besson, A. Papadopoulos and M. Troyanov IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society (EMS), Zürich, to appear 2014.
- [8] O. Hatori and L. Molnár, Isometries of the unitary groups and Thompson isometries of the spaces of invertible positive elements in  $C^*$ -algebras. *J. Math. Anal. Appl.* **409**(1), (2014), 158–167.
- [9] D. Hilbert, Über die gerade Linie als kürzeste Verbindung zweier Punkte. *Math. Ann.* **46** (1895), 91–96.
- [10] P. de la Harpe, On Hilbert's metric for simplices. *Geometric Group Theory, Vol. 1* (Sussex, 1991), London Math. Soc. Lecture Note Ser. 181, Cambridge Univ. Press, 1993, pp. 97–119.
- [11] B. Lemmens and R. Nussbaum, *Nonlinear Perron-Frobenius theory*. Cambridge Tracts in Mathematics, 189. Cambridge University Press, Cambridge, 2012.
- [12] B. Lemmens and R. Nussbaum, Birkhoff's version of Hilbert's metric and its applications in analysis. In G. Besson, A. Papadopoulos and M. Troyanov editors, Handbook of Hilbert geometry, IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society (EMS), Zürich, to appear 2014.
- [13] B. Lemmens and M. Roelands, Unique geodesics for Thompson's metric, *Ann. Inst. Fourier (Grenoble)*, to appear. ([arXiv:1305.6643](#))
- [14] B. Lemmens and C. Walsh, Isometries of polyhedral Hilbert geometries. *J. Topol. Anal.* **3**(2), (2011), 213–241.
- [15] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. II. Function spaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete, 97. Springer-Verlag, Berlin-New York, 1979.
- [16] L. Molnár, Thompson isometries of the space of invertible positive operators, *Proc. Amer. Math. Soc.* **137**(11) (2009), 3849–3859.

- [17] R.D. Nussbaum, Hilbert's projective metric and iterated nonlinear maps, *Mem. Amer. Math. Soc.* **75**, (1988).
- [18] A.C. Thompson, On certain contraction mappings in a partially ordered vector space. *Proc. Amer. Math. Soc.* **14**, (1963), 438–443.
- [19] C. Walsh, Gauge-reversing maps on cones, and Hilbert and Thompson isometries, [arXiv:1312.7871](#), 2013.

SCHOOL OF MATHEMATICS, STATISTICS & ACTUARIAL SCIENCE, CORNWALLIS BUILDING, UNIVERSITY OF KENT, CANTERBURY, KENT CT2 7NF, UK. TEL: +44–1227823651, FAX: +44–1227827932,

*E-mail address:* `B.Lemmens@kent.ac.uk`

SCHOOL OF MATHEMATICS, STATISTICS & ACTUARIAL SCIENCE, RUTHERFORD ANNEXE, UNIVERSITY OF KENT, CANTERBURY, KENT CT2 7NX, UK.

*E-mail address:* `mark.roelands@gmail.com`

SCHOOL OF MATHEMATICS, STATISTICS & ACTUARIAL SCIENCE, RUTHERFORD ANNEXE, UNIVERSITY OF KENT, CANTERBURY, KENT CT2 7NX, UK.

*E-mail address:* `M.Wortel@kent.ac.uk`